Dear Editor and Reviewer # 2:

This study provides a longer term soil moisture dataset (ChinaCropSM1km) for crop drylands across mainland of China. ChinaCropSM1km perform better than public product in both higher accuracy and more details (daily, more soil layers) by using machine learning technology. Such soil moisture dataset with higher resolutions is very valuable for the studies on crop model, yield estimation, and climate change impact assessment. Moreover, their methodology is robust, and their interesting results were well interpreted. The irrigation module is a novel way to improve highly moisture estimation. Therefore, I recommend it can be accepted after a minor revision.

We appreciate your insightful comments on our paper. The comments offered have been immensely helpful. We have responded to every question, indicating exactly how we addressed each concern or problem and describing the changes we have made. The revisions have been approved by all authors. The point-to-point responses to your comments are listed below in **blue**.

Comments and suggestions:

Point 1: There is a problem with the resolution. The ground observation data is point measurement data, how to match the resolution of 1km? Please explain this in the manuscript.

Response: Thank you a lot for the insightful suggestion. We have followed you to insert relevant contents into our manuscript (highlighted in "Track Changes", line 140~143, 185~186).

"We use the Extract Values to Points tool to extract the 1km resolution raster information of environmental (i.e. SP, RSD and GI) data to AMS point data, output point data attributes and save it in CSV format to obtain a data set of environmental factors through ArcGIS 10.5.".

"All these point samples are used to develop pointed-SM model, and then applied these pointed-models developed to inversely calculate the gridded-SM by inputting 1km-raster environmental variables.".

Point 2: Section 2.1. The authors pointed out that the study area is dominated by dryland crops (i.e.

wheat and maize) in China, how was the Chinacropland layer defined in Figure 1 according to the annual crop harvested area in mainland China from 2000 to 2015? please describe the details.

Response:

We remain the ChinaCropland location constant as several publications did similarly (Gervois et al., 2008; Ke et al., 2018). We proposed a new crop phenology-based crop mapping approach to generate a 1 km harvesting area dataset for three staple crops (i.e. rice, wheat, and maize) in China from 2000 to 2015 based on GLASS leaf area index (LAI) products (Luo et al., 2020a, b). Actually, we used the union of the annual harvested area dataset for maize and wheat as the China crop drylands maps.

Reference:

Gervois, S., Ciais, P., de Noblet-Ducoudré, N., Brisson, N., Vuichard, N., and Viovy, N.: Carbon and water balance of European croplands throughout the 20th century: CARBON BALANCE OF EUROPEAN CROPLANDS, Global Biogeochem. Cycles, 22, n/a-n/a, https://doi.org/10.1029/2007GB003018, 2008.

Ke, X., van Vliet, J., Zhou, T., Verburg, P. H., Zheng, W., and Liu, X.: Direct and indirect loss of natural habitat due to built-up area expansion: A model-based analysis for the city of Wuhan, China, Land Use Policy, 74, 231–239, https://doi.org/10.1016/j.landusepol.2017.12.048, 2018.

Luo, Y., Zhang, Z., Chen, Y., Li, Z., and Tao, F.: ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 2000–2015 based on leaf area index (LAI) products, Earth Syst. Sci. Data, 12, 197–214, https://doi.org/10.5194/essd-12-197-2020, 2020a.

Luo, Y., Zhang, Z., Li, Z., Chen, Y., Zhang, L., Cao, J., and Tao, F.: Identifying the spatiotemporal changes of annual harvesting areas for three staple crops in China by integrating multi-data sources, Environ. Res. Lett., 15, 074003, https://doi.org/10.1088/1748-9326/ab80f0, 2020b.

Point 3: In (1), the author judges the irrigation factors by comparing the observed soil moisture and the soil moisture evaluation index (SMI) according to the corresponding soil depth and phenology of crops. However, I notice that the SMI in Table 2 is a range, rather than an exact number. Please give reasonable explanation for this.

Response:

Actually, we use the minimum value of the SMI interval (i.e. an exact number threshold) to judge the irrigation factors considering the spatial differences in irrigated cropland. The irrigation factor (CIR) is assigned by 1 if the actual soil moisture is larger than the irrigation threshold. We used the minimum value to ensure that CIR were taken into account in all zones during forecasting SM. Using the minimum value might misclassify CIR, such as assigning "1" to no irrigation application, but such treatment is a compromise way before more detailed irrigation information is available. Moreover, we explained this limitation in the discussion section.

Point 4: In section 2.3.2, considering the new SM product has been derived by integrating the irrigation module into SM model, it is better to evaluate accuracy of the module (irrigation factor forecasting model) and supply such important information into new edition.

Response: In the Data and methods (Section 2.2.1), the accuracy of irrigation factor forecasting model has been provided in the revised manuscript (Line 214~220).

We evaluated our irrigation factor forecasting model results (**Table S4**) using the receiver operating characteristics (ROC) curve and their Area Under the Curve (AUC) (Fawcett, 2006). Also, we calculated UA (Eq. 7), PA (Eq. 8), and overall accuracy (Eq. 9) based on confusion matrices (**Table S3**) containing the percentages of the four possible outcomes of a model: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) (Fawcett, 2006).

$$PA = \frac{TP}{TP+FP} \tag{7}$$

$$UA = \frac{TP}{TP+FN} \tag{8}$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{9}$$

Reference:

Fawcett, T.: An introduction to ROC analysis, Pattern Recognition Letters, 27, 861–874, https://doi.org/10.1016/j.patrec.2005.10.010, 2006.

Table S3 Confusion matrix table in this study.

		Class	
		Irrigated	Non
Reference	Irrigated	TP	FN
	Non	FP	TN

Table S4 Confusion matrix of irrigated validation based on the test dataset. Prediction categories are columns while reference categories are rows.

ChinaCropSM1km	Class	Irrigated	Non	Total	Accuracy	PA	UA	AUC
$wheat_{0-10}$	Irrigated	1633	395	2028	0.85	0.82	0.81	0.84
	Non	365	2744	3109				
	Total	1998	3139					
$wheat_{10-20}$	Irrigated	1583	446	2029	0.84	0.81	0.78	0.83
	Non	365	2749	3114				
	Total	1948	3195					
$maize_{0-10}$	Irrigated	915	310	1225	0.86	0.85	0.75	0.84
	Non	167	2030	2197				
	Total	1082	2340					
$maize_{10-20}$	Irrigated	875	321	1196	0.86	0.83	0.73	0.83
	Non	175	2052	2227				
	Total	1050	2373					

Point 5: Some typos are found in manuscript, and check manuscript carefully and correct them. e.g. Line143: delete 'in China'.

Response: Thank you for your careful comments. We have modified all typos in the revised paper (Line 58, 100, 143, 524).

Line 58: r^2 -> R^2

Line 100: “accumulated precipitation for 10 days” -> “ante-accumulated precipitation over ten days”

Line 143: ~~in China~~

Line 524: “mode” -> “factor”

Point 6: Figure 2 should be improved. Currently, some labels are too vague to clearly identify.

Response: Many thanks for your advice. We have modified it in the revised paper.

Point 7: Please modify the line widths in Table 2.

Response: Many thanks for your careful check. We have modified it in the revised paper.

Point 8: Line257: insert blank between two words. ‘Figure8’ -> ‘Figure 8’.

Response: Thanks for your careful review. We have modified it in the revised paper. (Line 284).

Point 9: Figure S5 was not used in the main text, please cite it in main text or delete it from supplemental material.

Response: Thanks for your careful review. We have deleted it from supplemental material.